Московский авиационный институт

(национальный исследовательский университет)

Институт № 8 «Компьютерные науки и прикладная математика»

Кафедра 804 «Теория вероятностей и компьютерное моделирование»

**Курсовая работа**

**По дисциплине «Теория вероятности и математическая статистика»**

**По теме: «Метод наименьших квадратов»**

Выполнил студент группы М8О-305Б-20

Попова Наталья Сергеевна

Преподаватель: Ибрагимов Данис Наилевич

Оценка:

Дата: 26.12.22

Москва, 2022

**Описание модели**

Модель полезного сигнала имеет вид:

Рассматривается модель наблюдений

где независимые и одинаково распределённые случайные величины.

**Моделирование данных**

Смоделировать два набора наблюдений на основе модели (2) для следующих случаев:

|  |  |
| --- | --- |
| 1 случай | 2 случай |
|  |  |
|  | |

Параметры задания определяются номером варианта:

|  |
| --- |
| Вариант 16 |
|  |

**Задание**

Для обоих случаев выполнить по очереди следующие задания.

**1.** Подобрать порядок многочлена в модели (1), используя критерий Фишера, и вычислить оценки неизвестных параметров методом наименьших квадратов.

**2.** В предположении нормальности ошибок построить доверительные интервалы уровней

надёжности и для параметров .

**3.** В предположении нормальности ошибок построить доверительные интервалы уровней

надёжности и для полезного сигнала (1).

**4.** Представить графически

• истинный полезный сигнал,

• набор наблюдений,

• оценку полезного сигнала, полученную в шаге 1,

• доверительные интервалы полезного сигнала, полученные в шаге 3.

**5.** По остаткам регрессии построить оценку плотности распределения случайной ошибки наблюдения в виде гистограммы.

**6.** Вычислить оценку дисперсии случайной ошибки.

**7.** По остаткам регрессии с помощью -критерия Пирсона на уровне значимости 0.05 проверить гипотезу о том, что закон распределения ошибки наблюдения является нормальным.

**1 случай**

**Задание 1.**

Для оценки порядка многочлена модели полезного сигнала проверяются гипотезы при *,*  т.е. рассматриваемая модель наблюдения будет выглядеть:  *–* принимается. В таком случае .

Для проверки гипотез нужно найти статистику, которую можно посчитать по формуле:

где

– старший коэффициент,

*–* оценка вектора ошибок,

*n –* объем выборки,

- элемент *j-*ой строки *j-*ого столбца матрицы ,

*j* – степень многочлена, для которой проверяется гипотеза.

Гипотеза принимается если т.е. если

*Метод наименьших квадратов (МНК)* — математический метод, применяемый для решения различных задач, основанный на минимизации суммы квадратов отклонений некоторых функций от экспериментальных входных данных. Формула:

где

– оцениваемые параметры

*X* – матрица элементов

*Y* – матрица наблюдений

Предположим, что:

m = 1: T() = 0.8966 < = 2.0244 => гипотеза принимается

m = 2: T() = 94.4081 > = 2.0262 => гипотеза отвергается

m = 3: T() = 2.1916 > = 2.0281 => гипотеза отвергается

m = 4: T() = 0.5674 < = 2.0301 => гипотеза принимается

Следовательно,

Оценки неизвестных параметров:

**Задание 2.**

*Доверительный интервал* —интервал, который покрывает неизвестный параметр с заданной надёжностью. Общая формула для нашего случая выглядит следующим образом:

Выразим оцениваемую величину и получим следующую формулу:

Где

– уровень надежности,

- элемент k-ой строки k-ого столбца матрицы ,

– оценка вектора ошибок, полученная в задании №1,

*n –* объем выборки,

*m –* степень многочлена,

– оценки параметров .

Доверительные интервалы для уровня надежности для

: (15.2680 << 15.3892)

: (3.5634 < < 3.6845)

: (-6.0298 < < -5.9087)

: (-0.1260 < < - 0.0049)

Доверительные интервалы для уровня надежности для

: (15.2474 << 15.4098)

: (3.5427 < < 3.7051)

: (-6.0504< < - 5.8880)

: (-0.1467< < 0.0158)

**Задание 3-4.**

Доверительный интервал для полезного сигнала *Y* можно вычислить по следующей формуле:

где

– интервал надежности,

- элементы на главной диагонали матрицы ,

– оценка модели полезного сигнала,

– оценка вектора ошибок, полученная в задании №1,

*n –* объем выборки,

*m –* степень многочлена.

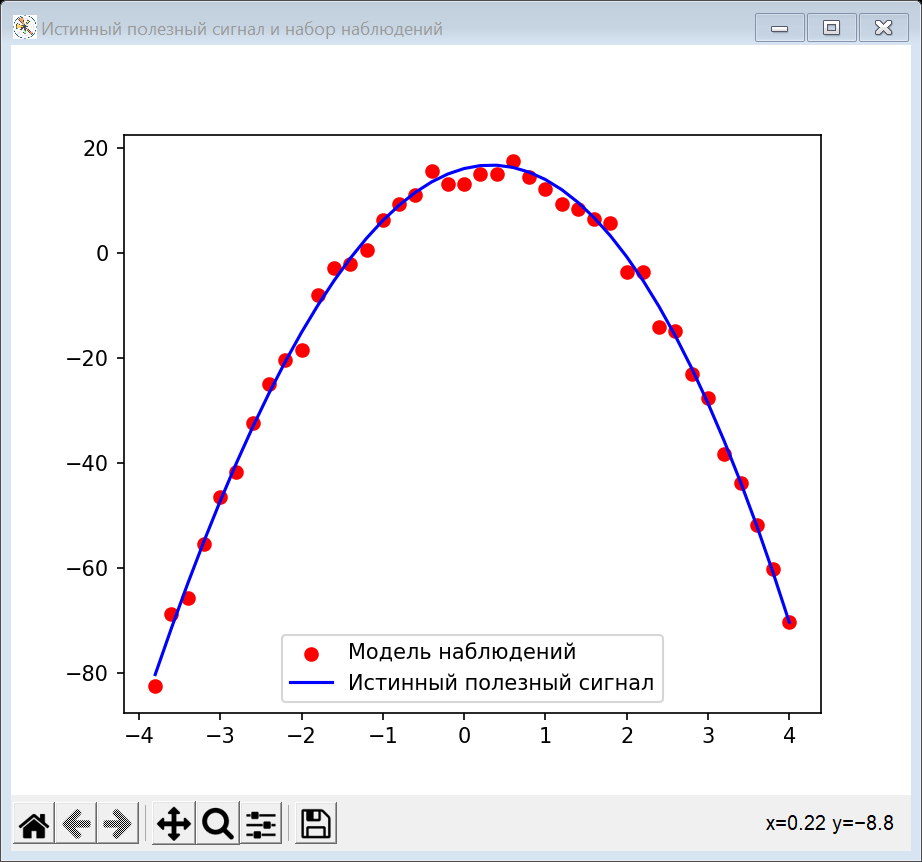


Рисунок - Модель наблюдений и истинный полезный сигнал

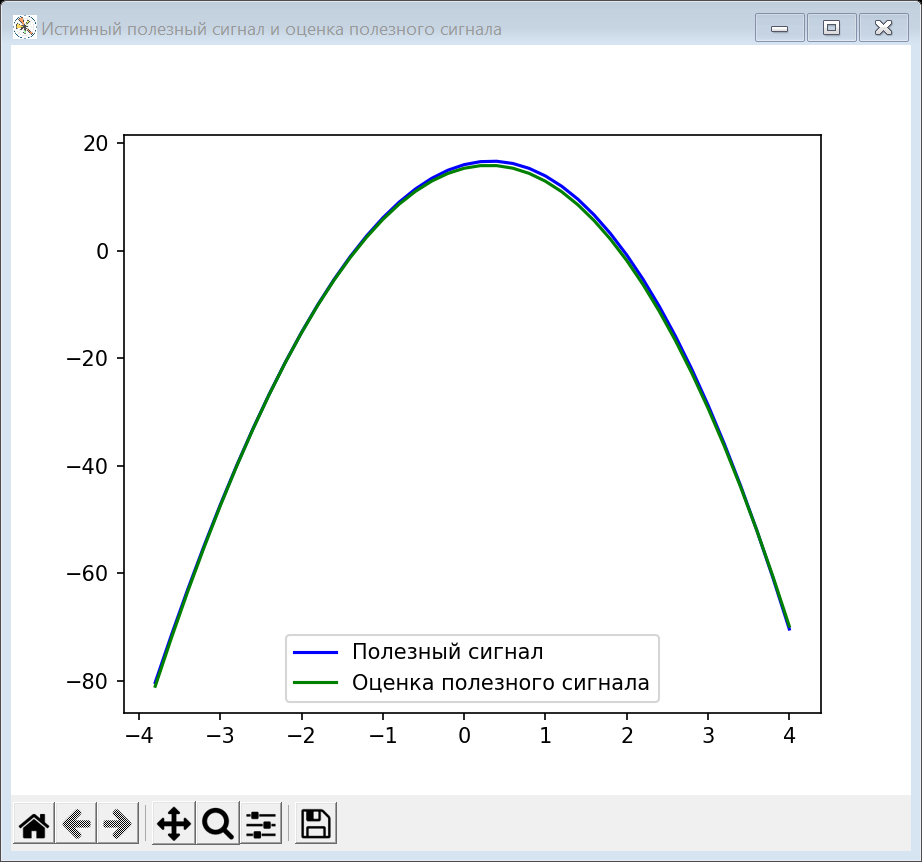


Рисунок - Истинный полезный сигнал и оценка полезного сигнала

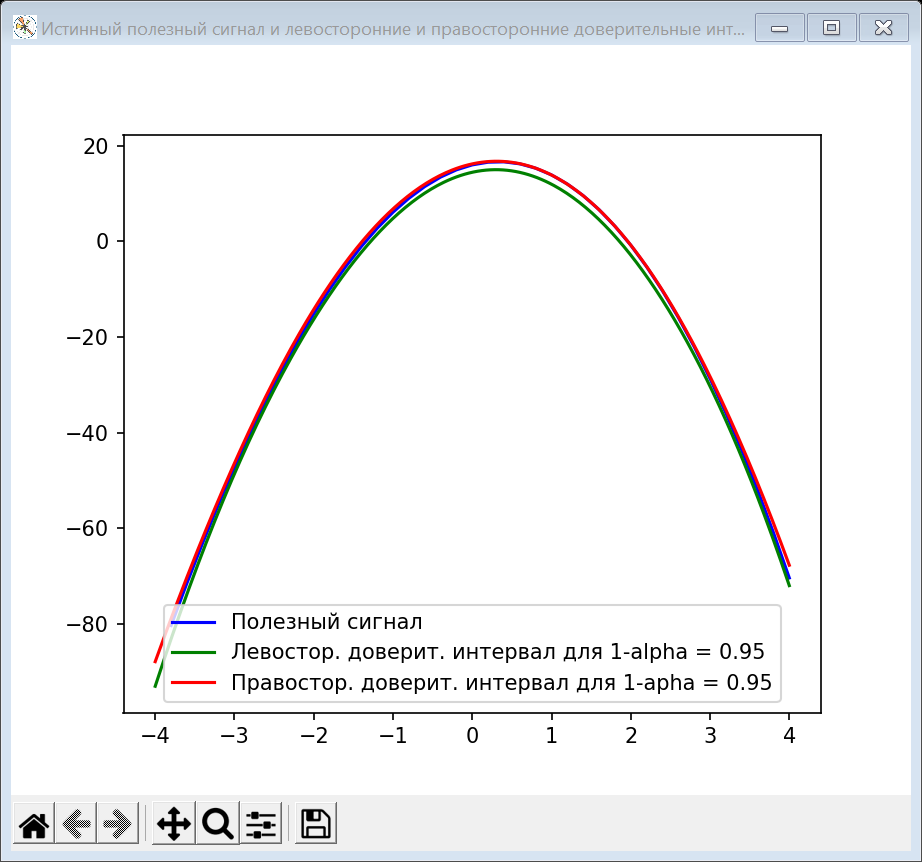


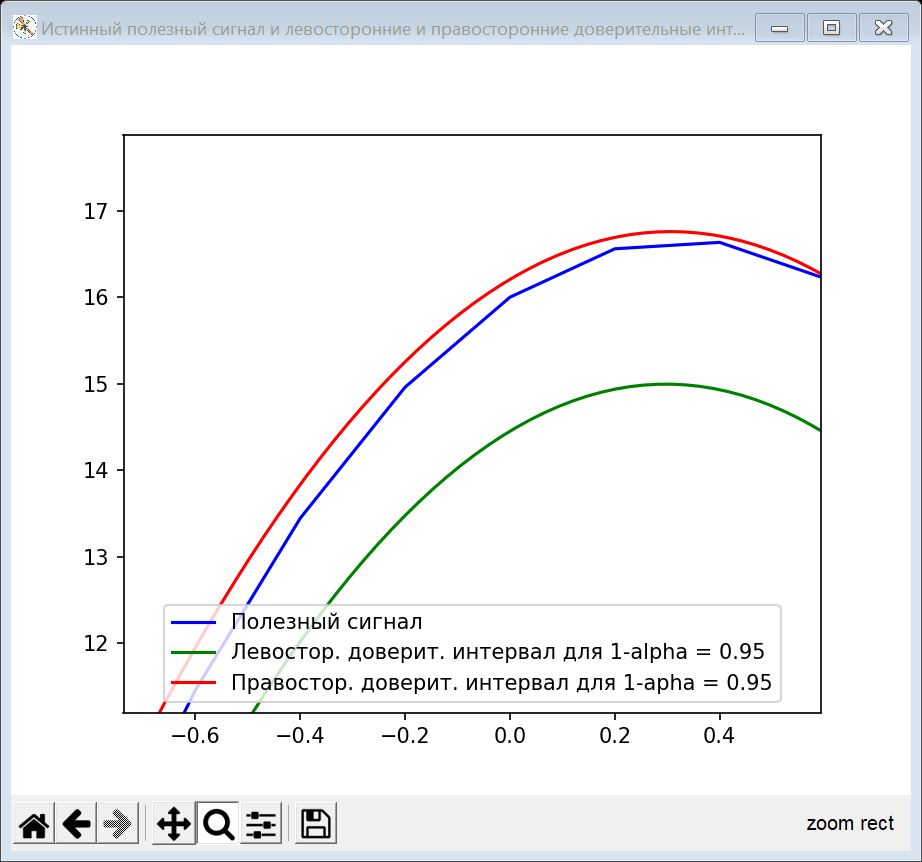
Рисунок - Истинный полезный сигнал и левосторонние и правосторонние доверительные интервалы для а = 0.95

Рисунок - Истинный полезный сигнал и левосторонние и правосторонние доверительные интервалы для а = 0.95

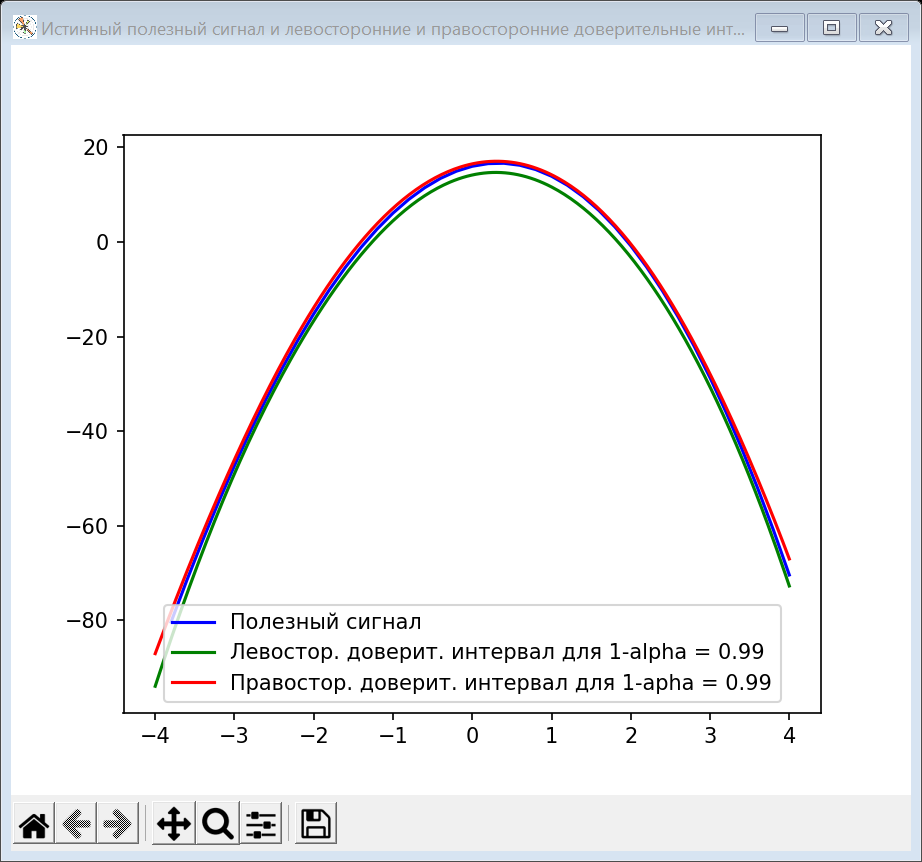


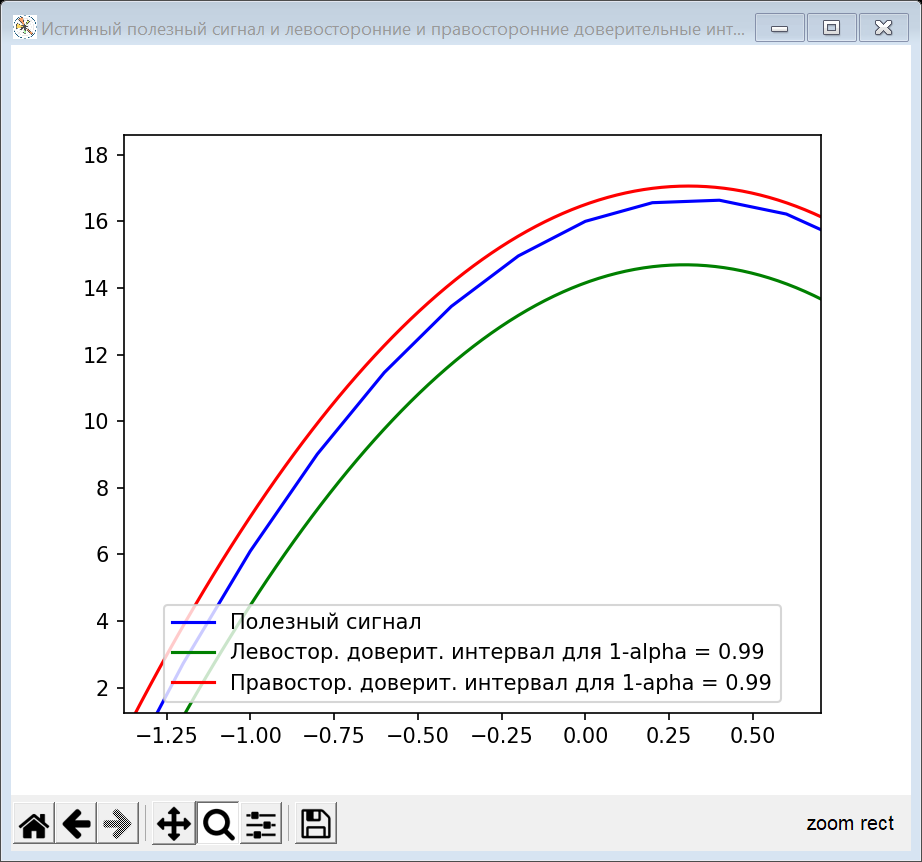
Рисунок - Истинный полезный сигнал и левосторонние и правосторонние доверительные интервалы для а = 0.99

Рисунок - Истинный полезный сигнал и левосторонние и правосторонние доверительные интервалы для а = 0.99

**Задание 5.**

Чтобы построить *гистограмму* плотности распределения ошибки наблюдения, разобьем отрезок на равных промежутков. Воспользуемся следующей формулой для вычисления оценки вероятности нахождения значения в промежутке (*k*-ая точка разбиения ):

где

n – объем выборки,

– кол-во значений , попавших в *i*-ый промежуток разбиения , .

Чтобы получить нормированную гистограмму, воспользуемся формулой

Значение *l* было выбрано равным 7. Значения соответствуют вероятности попадания элементов *E* в промежутки и соответственно, но т.к. все значения *E* лежат в отрезке ⇒

Также поверх гистограммы выведем плотность распределения

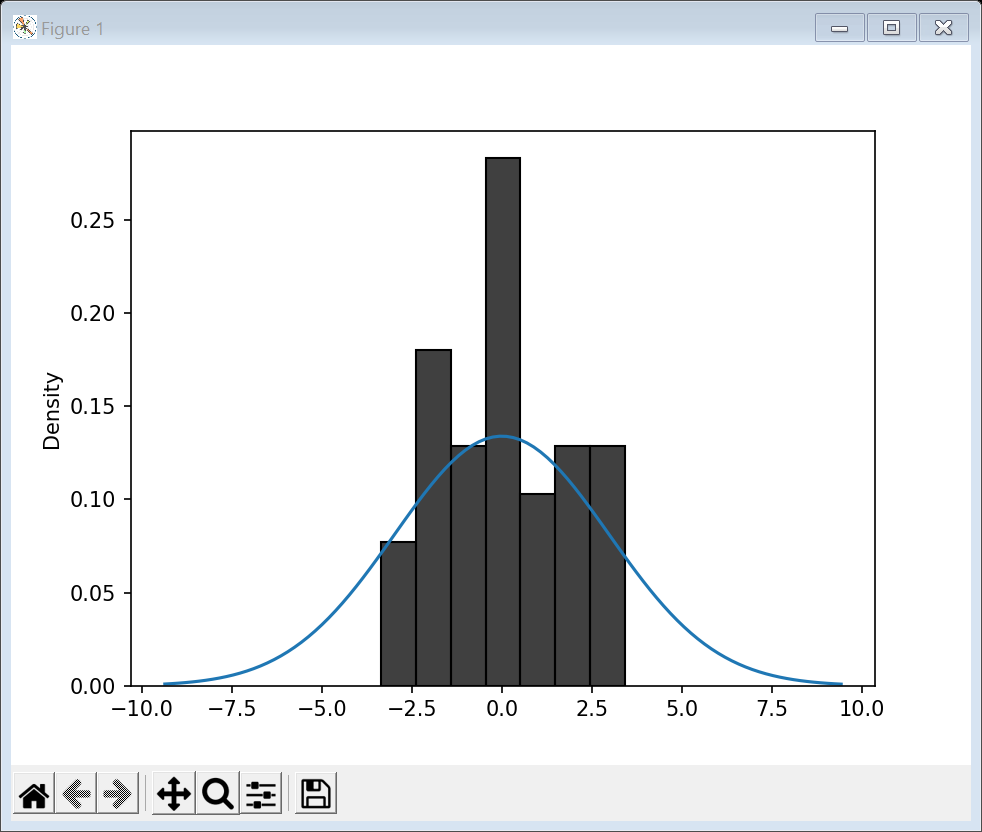
****

Рисунок - Оценка плотности распределения ошибки наблюдения и плотность распределения N(0,σ^2)

**Задание 6.**

Оценка *дисперсии* вычисляется по формуле:

где:

*n* – объем выборки,

– оценка вектора ошибок

Оценка дисперсии случайной ошибки: = 3.4400

**Задание 7.**

*-критерия Пирсона* – наиболее часто употребляемый критерий для [проверки гипотезы](https://ru.wikipedia.org/wiki/%D0%9F%D1%80%D0%BE%D0%B2%D0%B5%D1%80%D0%BA%D0%B0_%D1%81%D1%82%D0%B0%D1%82%D0%B8%D1%81%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D1%85_%D0%B3%D0%B8%D0%BF%D0%BE%D1%82%D0%B5%D0%B7) о принадлежности наблюдаемой выборки  {\displaystyle x\_{1},x\_{2},...,x\_{n}}объёмом *n* {\displaystyle n}некоторому теоретическому [закону распределения](https://ru.wikipedia.org/wiki/%D0%A0%D0%B0%D1%81%D0%BF%D1%80%D0%B5%D0%B4%D0%B5%D0%BB%D0%B5%D0%BD%D0%B8%D0%B5_%D0%B2%D0%B5%D1%80%D0%BE%D1%8F%D1%82%D0%BD%D0%BE%D1%81%D1%82%D0%B5%D0%B9){\displaystyle F(x,\theta )}. В нашем случае будет проверяться гипотеза о принадлежности вектора ошибок нормальному закону распределения:

Для проверки гипотезы нужно найти статистику, которую можно посчитать по формуле

где

*n* – объем выборки,

*l* – число столбцов гистограммы, построенной в задании 5.

–вероятность нахождения значений вектора ошибок в *i*-ом интервале, вычисленная в задании 5,

где

функция Лапласса,

– *k*-ая точка разбиения значений вектора ошибок.

Гипотеза будет принята нами, если , где *l* – число столбцов гистограммы, т.е. если . В противном случае гипотеза будет отвергнута.

= 10.2535

10.2535 < , значит гипотеза о том, что распределение нормальное, принимается.

**2 случай**

**Задание 1.**

m = 1: T() = 1.5103 < = 2.0244 => гипотеза принимается

m = 2: T() = 33.9652 > = 2.0262 => гипотеза отвергается

m = 3: T() = 0.9700 < = 2.0281 => гипотеза принимается

Следовательно,

Оценки неизвестных параметров:

**Задание 2.**

Доверительные интервалы для уровня надежности для

: (15.4297 << 16.1521)

: (3.9589 < < 4.6813)

: (-6.4161 < < -5.6937)

Доверительные интервалы для уровня надежности для

: (15.3068 << 16.2750)

: (3.8360 < < 4.8042)

: (-6.5390 < < -5.5708)

**Задание 3-4.**

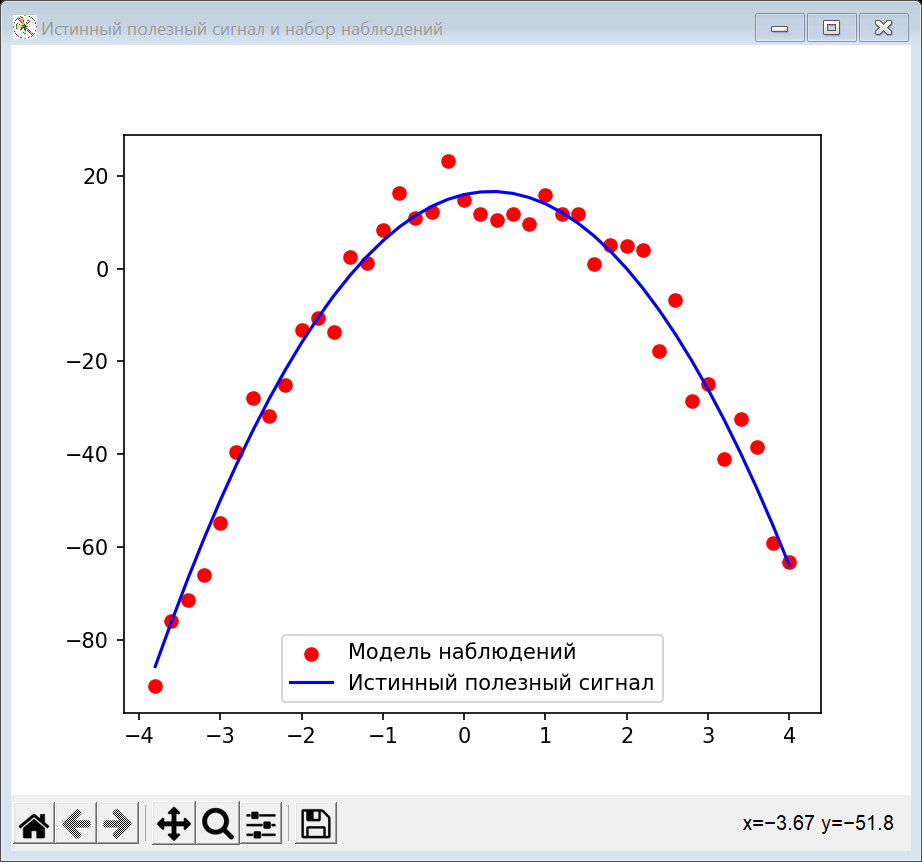


Рисунок - Истинный полезный сигнал и набор наблюдений

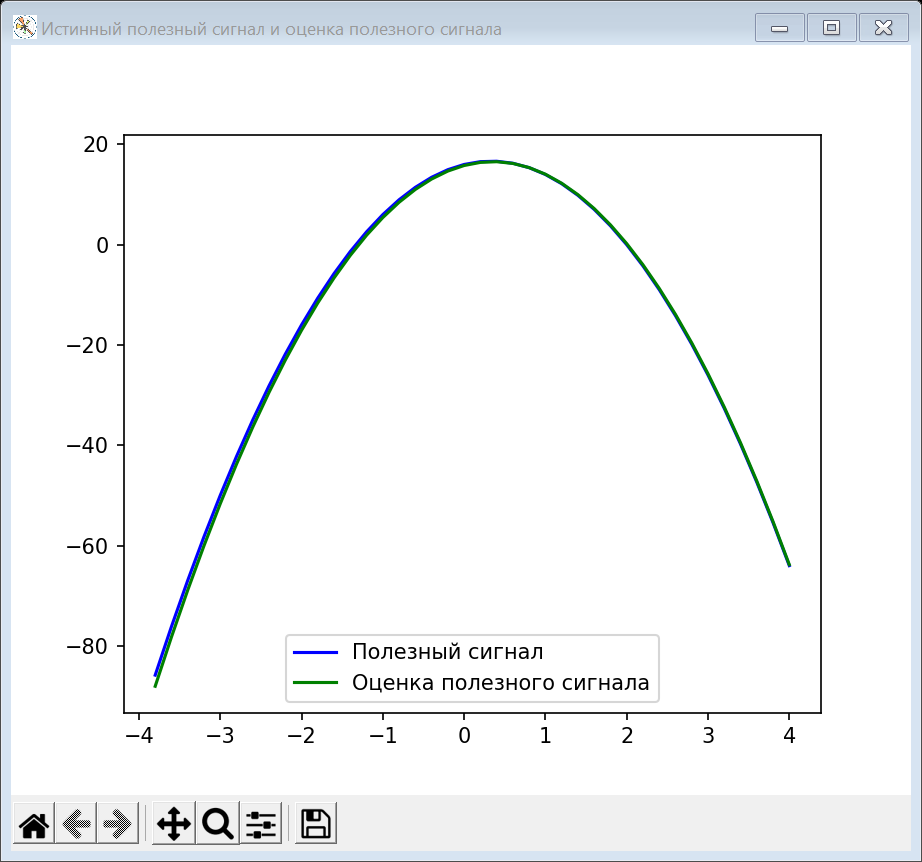


Рисунок - Истинный полезный сигнал и оценка полезного сигнала

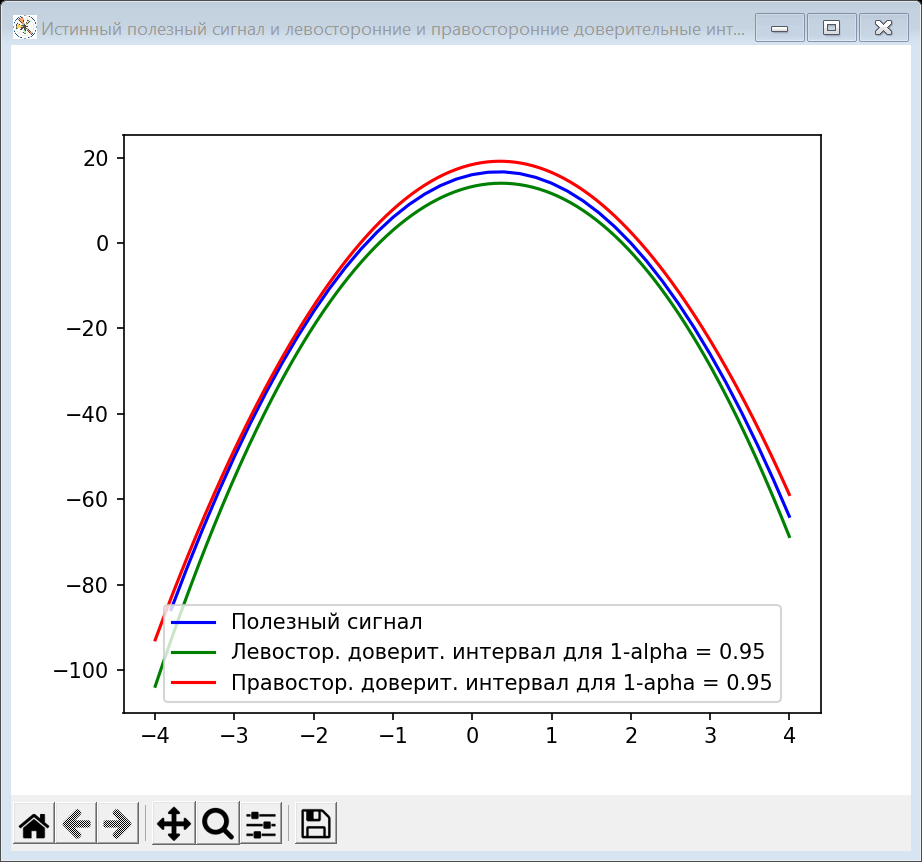


Рисунок - Истинный полезный сигнал и левосторонние и правосторонние доверительные интервалы а = 0.95

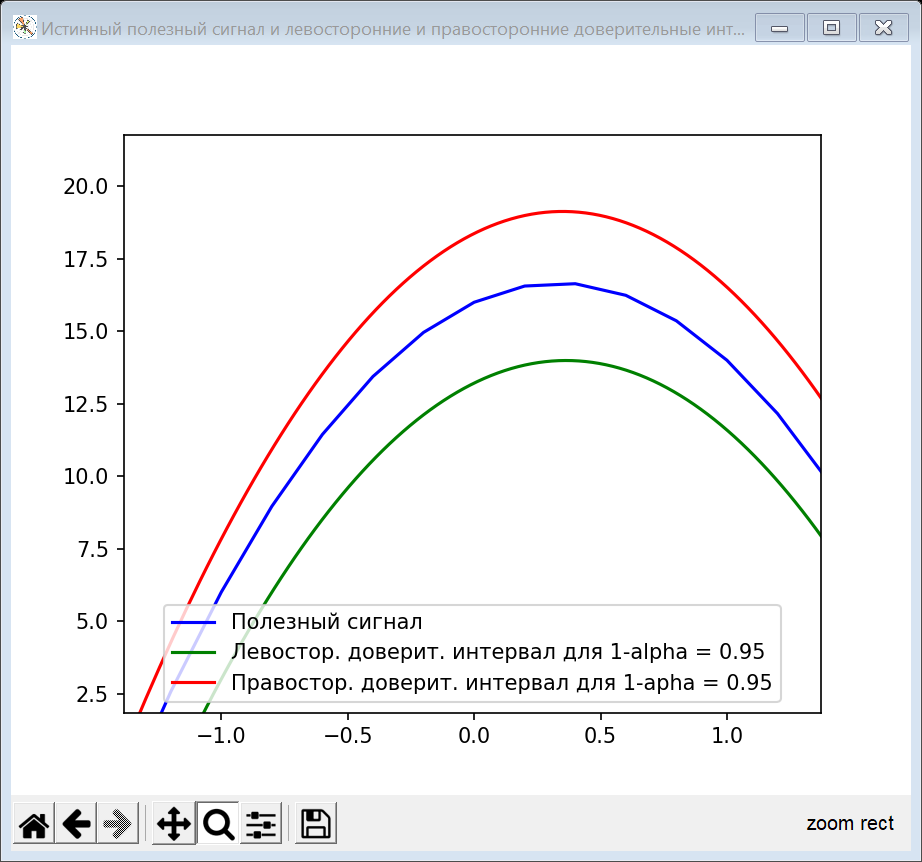


Рисунок - Истинный полезный сигнал и левосторонние и правосторонние доверительные интервалы а = 0.95

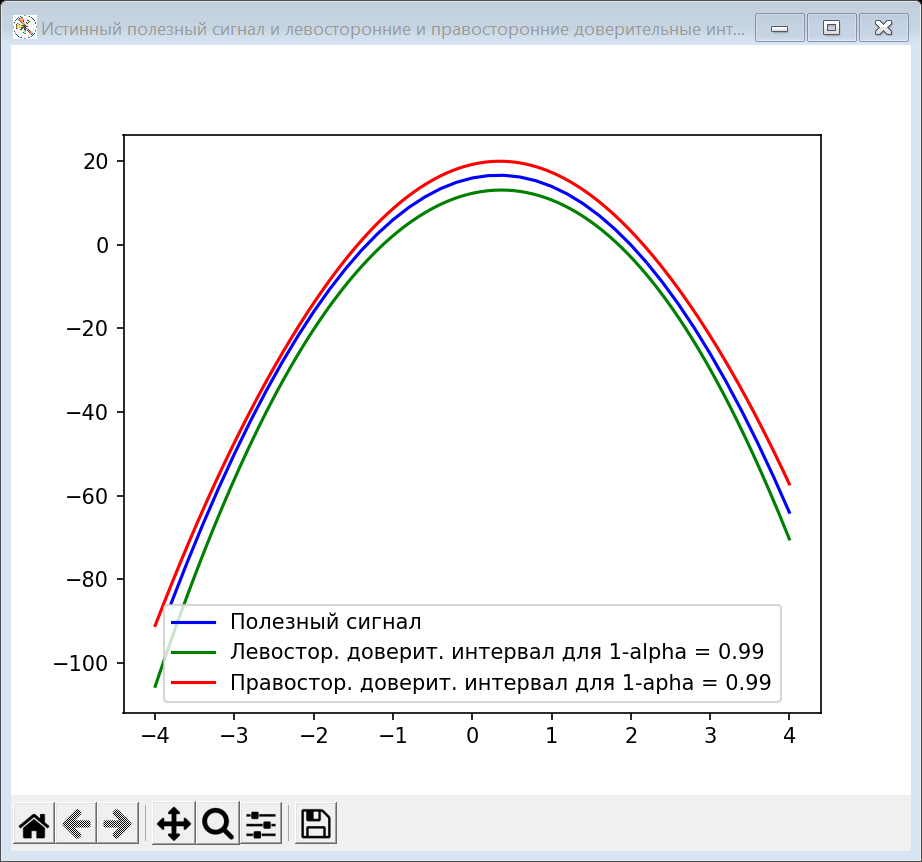


Рисунок - Истинный полезный сигнал и левосторонние и правосторонние доверительные интервалы а = 0.99

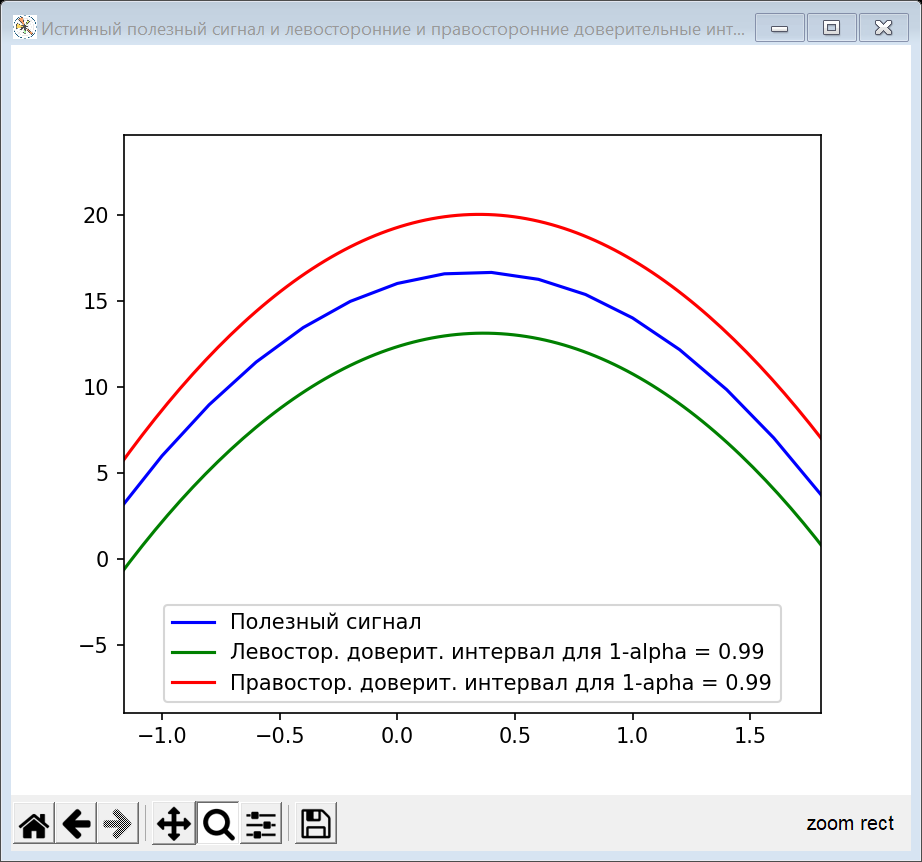


Рисунок - Истинный полезный сигнал и левосторонние и правосторонние доверительные интервалы а = 0.99

**Задание 5.**

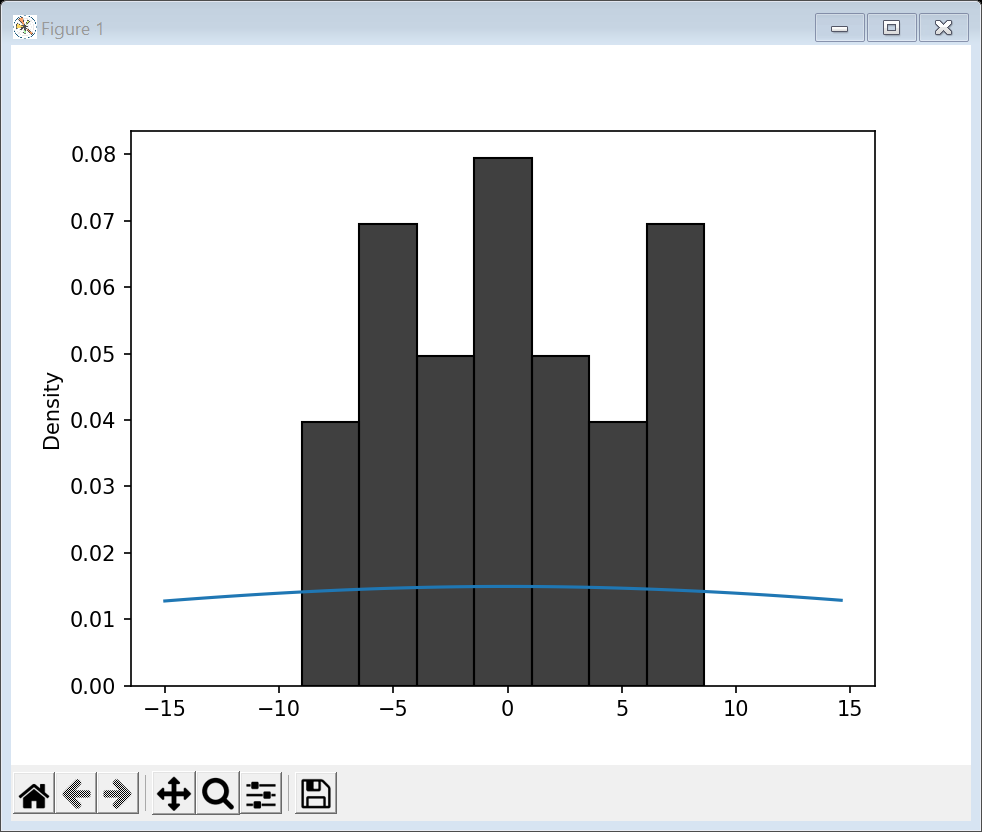
****

Рисунок - Оценка плотности распределения ошибки наблюдения и плотность распределения N(0,σ^2)

**Задание 6.**

Оценка дисперсии случайной ошибки: = 26.6733

**Задание 7.**

= 92.3026

92.3026 > , значит гипотеза о том, что распределение нормальное, отвергается